Improved Combined Radar/Radiometer Rain Profiling

Z. S. Haddad and J. P. Meagher
Jet Propulsion Laboratory
California Institute of Technology, Pasadena, CA 91109 USA
(818)354-1218/(818)393-5285/zsh@jpl.nasa.gov

Abstract

The current TRMM combined radar/radiometer profiling algorithm compensates for the known shortcomings of each instrument by exploiting the strengths of the other. It turns out that the strengths/weaknesses of the radiometer measurements are not as they seemed before TRMM. Specifically, the current algorithm presumes that the radiances at the various frequencies are approximately independent (to compute probability weights). It turns out there are strong correlations between the channels. New estimates of the conditional covariance of the radiances given the rain should make the probability weights more realistic. The second problem stems from the representativity of the TRMM cloud database. It turns out that the database is significantly off, especially at higher frequencies. To reduce the rain over-estimation produced by this discrepancy, the database was re-sampled and mean rain-radiances relations were re-derived. The third problem is the lack of any ice estimates. The reason was the large number of unknown variables involved. A principal component analysis has revealed that the frozen hydrometeor profiles can be approximated by a single variable each for ice, snow, and graupel. A straightforward method is currently being implemented to estimate these additional variables, and to include them in the output structure of the algorithm. Linear formulas will enable users to reconstruct the corresponding graupel/snow/ice profiles, and estimate the corresponding latent heating.

Introduction

The "day-1" version of the combined algorithm that was operational at launch was deliberately kept simple to reduce the opportunities for serious errors. We followed the procedure described in ([3]): **a)** grouping the radar beams according to the TMI beams within which they fall, we consider that several values for the rain drop-size-distribution (DSD) normalized mean diameter D'' are possible (see [4] for the precise definition of the DSD parametrization used); **b)** for each possible value of D'', D''-specific Z–R and k–R relations are used to invert each radar reflectivity profiles within the TMI-delimited group into a possible (D''-dependent) rain-rate profile, using an inversion that is very close to the one used by the radar-only profiling algorithm 2A25; **c)** then, for each value of D'', the algorithm calculates from the rain profiles a predicited 10.7GHz V-pol radiance $T(D'')$, the rain-radiance relation having been derived off-line by taking an average over numerous cloud simulations; **d)** finally, the difference between the predictions and the measured radiance τ is used to calculate a weight $w(D'') = \exp(-0.5\,(T(D'')-\tau)^2/\sigma_\tau^2)$ which quantifies how well (or badly) the radiance corresponding to the radar-derived rain-rates for each D'' fit the observed radiance (the amount of "slack" σ_τ that must be allowed when considering the closeness of the candidate radiances to the measurement is essentially the τ-dependent error bar from the same off-line cloud simulations that produced the average rain-radiance relation). The combined estimate of the rain-rate profile is the average of the D''-dependent radar-derived profiles weighted by the TMI-derived $w(D'')$'s.

This "day-1" algorithm is based on the Bayesian mathematical approach described in ([3]). As such, it minimizes the variance between the estimated rain and the actual rain, given the measured radar reflectivities and the measured 10.7GHz V-pol radiance (and given the admittedly simplified physical model for the dependence of the observations on the rain). Some time after the launch of TRMM, we derived and implemented an improved version of this algorithm, one that could take into account the "tall vector" consisting of all the TMI observations rather than just the 10.7GHz radiance. To that end, step **c)** in the day-1 algorithm was modified so that the "tall-vector" algorithm computes predicted 19GHz, 22GHz and 37GHZ radiances from each radar-derived D''-dependent rain profile. These rain-radiances relations were derived off-line by calculating conditional means using the TMI cloud simulations database (the one used for the passive microwave profiling algorithm 2A12), after having classified the samples in the database according to the integrated ice content – three levels of ice are used, and the appropriate relation is selected according to the value of the 85GHz TMI radiance compared with two thresholds (a value below the low threshold forces the algorithm to use the "high-ice" conditional mean rain-radiances relations, while a value above the high threshold selects the "low-ice" relations). In addition, step **d)** was modified to calculate the weights according to the total differences in all TMI channels. Thus the weights in the "tall-vector" algorithm are given by

$$w(D'') = \exp(-0.5 \sum_i (T_i(D'') - \tau_i)^2 / \sigma_{\tau_i}^2), \qquad (1)$$

in which the σ_{τ_i}'s used are admittedly somewhat heuristic estimates of the τ_i-dependent error bars from the 2A12 cloud simulation database.

Comparisons between the "day-1" surface rain estimates and the surface rain estimated by 2A25 over several orbits show good agreement. Only in the case of stratiform rain with significant radar attenuation does a bias appear between the two algorithms.

While the comparisons above (as well as time- and space-average comparisons) give a favorable impression of the performance of the combined algorithm, we have identified three significant weaknesses in the algorithm. These are listed by order of importance and described in the next section.

Problems and Improvements

The use of equation (1) to compute the weights to assign to the various DSD parameters in contention presumes that the various TMI measurements are independent. Lacking any quantitative information about the possible correlations between the brightness temperatures at the different frequencies and polarisations, this assumption of independence is reasonable and does appear to have produced an algorithm whose estimates are not flagrantly biased. However, our analysis of the joint behavior of the radiances in the 2A12 cloud simulations database has revealed very strong correlations between the various channels ([1]).

Our results in ([1]) imply that (1) should be replaced by the gaussian distribution with the appropriate (non-diagonal) conditional covariance matrix, computed for each ice category.

In addition, the calculation of the mean radiances T_i from the 2A12 cloud simulations database, in step **c)**, assumes that the database is representative of tropical rain. As a comparison of the marginal histograms of the database radiances and of the radiances actually measured by TMI shows, the database represents the actual brightness temperatures at 10.7GHz much better than the ones at 37GHz. This mismatch between the database and the actual measurements is quite likely to produce a bias, although the effect would be less pronounced in the "day-1" algorithm since the discrepancy between database and measurements is not very pronounced at 10.7GHz.

In order to obtain mean rain-radiances relations and corresponding conditional covariance matrices that faithlfully reflect the behavior of tropical rain and its microwave signature, it is necessary to start with a database which accurately mimics the joint statistics of the observations. To do so, starting with the current 2A12 cloud simulations database, we intend to sub-sample the simulated profiles using the extensive archive of observations which the TMI algorithm 1B11 has been producing over the last two years (this is the data that we used to produce the histograms in figures 14 and 16).

More specifically, we analyzed the covariance from our extensive set of granules, then found the formulas for the two eigenvectors T_1' and T_2' of the 9×9 covariance matrix with the highest eigenvalues; since these vectors are necessarily uncorrelated and therefore, to first order, essentially independent, we then partitioned the joint values into 16×14 "bins", of about 10 percentiles each. Given this partitioning scheme, for each sample in the original database, we can compute its T_1' and T_2' (linear combinations of the sample profile's brightness temperatures), and then select a proportion from each bin according to the sample proportions ($\simeq$ 10%) computed for the representative granules. The selected simulated profiles are not absolutely representative of all tropical rain, but they reproduce the joint behavior of the radiances in the raw TMI measurements, and as such form a database that is far more realistic than the original.

Results

Using the sub-sampled "improved" database obtained as described above, we derived average rain-radiances relations to replace the "naïve" ones currently implemented in 2B31, and thus avoid having to refer to any database in real time.

To that end, we used the result ([1]) that the "tall" vector of radiances can be quite adequately described using the two eigenvectors T_1' and T_2' (the first 3 eigenvalues for the covariance matrix of the brightness temperatures in the current 2A12 database are, in decreasing order, $5652 > 965 > 314$, while for our archive of 1B11 granules they are $1711 > 275 > 95$; in both cases, the third eigenvalue is already better than an order of magnitude below the first). The problem of estimating the radiances from a profile vector of rain-rates can therefore be reduced to estimating the corresponding values of T_1' and T_2'. Since it is operationally very desirable to use a subset of all the rain-rates to estimate T_i', $i = 1, 2$ (or, in effect, "distill" the information contained in the database into two parametrized functional relations, we looked for the "optimal" linear combination of the rain-rates that will "best" estimate each T_i'. Because the problem of finding the best relation between T_i' and a combination R' of the vertical rain-rates R_j is a priori non-linear, we modified it slightly by trying to maximize the correlation's numerator $E\{T_i' \cdot R'\}$ keeping $E\{R'^2\}$ constant. This in effect minimizes the scatter between R' and T_i'. Once the coefficients of R' were found, the mean and variances of T_i' given R' were easily obtained, along with the forward rain-radiances relation and its uncertainty.

We have tested this procedure on its exact mirror image, namely obtaining mean radiances-rain relations and the associated conditional covariance, using a (theoretically interesting) database of nadir passive microwave observations. In that case (taken from [1]), the first rain eigenvector (the equivalent of T_1' above) turns out to be simply the vertically-averaged rain-rate, and the optimal linear combination of radiances that correlates best with it (the equvalent of R' above) turns out to be $0.42T_{10.7H} + 0.8T_{19.3H} - 0.1T_{21.3V} - 0.06T_{37H} + 0.1T_{85.5H}$ The components of the vector reconstructed from the single "optimal" relation are encouragingly close to the exact values. The fact that the approach produces satisfactory results in the case of estimating the rain from the radiances leads us to believe that it will work very well indeed in the converse case of interest to us here, since we are then starting with the richer set of information (the rain profile) and looking to predict a small number of corresponding observations (the radiances).

As to the problem of estimating graupel, ice and snow profiles, along with a rain profile, the main obstacle is the large number of unknown variables involved. We used a principal component analysis to determine just how many variables were essential to describe a rain profile. A detailed analysis of the rain profiles simulated in the TRMM passive microwave cloud database showed that the set of variables $(R_1, \cdots, R_8)$, representing the rain rate in the eight layers of the atmosphere extending from the surface up to a height of 4.5 km, can be replaced with eight linear combinations $(R_1', \cdots, R_8')$ whose variances decrease so rapidly that the variance of R_4' is already an order of magnitude smaller than the variance of R_1' ([1]). This means that one can adequately describe the same rain profile using R_1' and R_2' and the archival mean values over all rain profiles in the database of the remaining six variables $(R_3', \cdots, R_8')$. A similar analysis of the rain profiles independently estimated by the TRMM radar rain-profiling algorithm confirmed this finding. It is interesting to note that in both cases the first new variable R_1' was the average of the rain rate over the column and the second was the difference between the average rain rates above and below 2.5 km. The fact that the measurements independently confirmed the result obtained from the simulations shows that this reduction in the number of variables required to describe the rain is not an artifact of the database and is therefore a potentially very useful tool.

Indeed, it turned out that such a reduction can be obtained not just for the rain profiles but also and in a more dramatic way for the graupel, snow and ice profiles. In fact, a "test" principal component analysis performed on a simulated squall line showed that the first variable G_1' for the graupel profiles is a slightly unevenly-weighted average of the graupel between 3 and 13.5 km (with a peak weight for the layer at 5 km), the second variable G_2' is the difference between the average graupel below and above 6.5 km, and the ratio of the variance of G_2' to G_1' is 0.075, or more than an order of magnitude. Similarly, this test case revealed that the first variable S_1' for the snow profiles is a slightly unevenly-weighted average of the snow between 5 and 14 km (with a peak weight for the layer at 9 km), and the first variable I_1' for the ice profiles is a slightly unevenly-weighted average of the ice between 8.5 and 15.5 km (with a peak weight for the layer at 12 km), the second eigenvariables in both cases having significantly lower variance.

These preliminary results show that knowledge of the four principal component variables G_1', G_2' and S_1', and I_1' should allow an adequate reconstruction of the corresponding hydrometeor profiles.

Acknowledgments

This work was performed by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA).

References

[1] D. Coppens, Z.S. Haddad and E. Im: "Estimating the uncertainty in passive microwave rain retrievals", *submitted to J. Atmos. Ocean. Tech.*, 1999.

[2] D. Coppens and Z.S. Haddad: "Effects of drop size variability on microwave brightness temperature calculations", *submitted to J. Geophys. Res.*, 1999.

[3] Z.S. Haddad, E.A. Smith, C.D. Kummerow, T. Iguchi, M.R. Farrar, S.L. Durden, M. Alves and W.S. Olson: "The TRMM 'Day-1' radar/radiometer combined rain-profiling algorithm", *J. Met. Soc. Japan*, vol. 75 no. 4, pp 799-809, 1997.

[4] Z.S. Haddad, D.A. Short, S.L. Durden, E. Im, S. Hensley, M.B. Grable and R.A. Black: "A New Parametrization of the rain drop size distribution", *I.E.E.E. Trans. Geosci. Remote Sensing*, vol. 35, no. 3, pp 532-539, 1997.